AD_____

Award Number: DAMD17-02-1-0597

TITLE: Retroelements and Genetic Instability in Breast Cancer

PRINCIPAL INVESTIGATOR: Victoria Perepelitsa Belancio
Prescott L. Deininger

CONTRACTING ORGANIZATION: Tulane University Health Sciences Center
New Orleans, LA 70112-2699

REPORT DATE: April 2005

TYPE OF REPORT: Annual Summary

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	April 2005	Annual Summary (1 Apr 02 - 31 Mar 05)

4. TITLE AND SUBTITLE
Retroelements and Genetic Instability in Breast Cancer

5. FUNDING NUMBERS
DAMD17-02-1-0597

6. AUTHOR(S)
Victoria Perepelitsa Belancio
Prescott L. Deininger

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Tulane University Health Sciences Center
New Orleans, LA 70112-2699

E-Mail: vperepe@tulane.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 Words)

LINE1 is a mammalian retroelement that contributes to genomic instability. The full extent of LINE-1 mobility in somatic tissues and tumors is not known. L1 expression is extremely low in most differentiated cells except for testis, but it is significantly elevated in breast malignancies. This suggests that posttranscriptional mechanisms are involved in limitation of L1 expression. We demonstrated that the use of the polyA sites located within the L1.3 genome limits the amount of full-length L1.3 mRNAs. L1 polyA signals can be functional when fragments of L1.3 element are inserted into 3' UTRs of genes. This unique attenuation mechanism helps to minimize the rate of L1 retrotransposition, but may also increase the negative impact of these events on the genome after their insertion. Human EST database searches suggest that the polyA signals may also play a role in regulation of L1 expression in a tissue and/or tumor specific manner with breast cancer tissues supporting the least efficient L1 polyadenylation. The EST data are strengthened by significant differences in the L1 RNA profiles between transiently transfected breast cancer and nonbreast cancer cell lines. These observations suggest a potential global change in the mechanism of polyadenylation process upon malignant transformation of mammary gland.

14. SUBJECT TERMS
Regulation of gene expression, LINE1, polyadenylation, genomic instability, tissue specific expression

15. NUMBER OF PAGES
23

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

Introduction.

LINE1 is the most abundant active human retroelement, and it contributes to genomic instability. The full extent of LINE-1 mobility in somatic tissues and tumors is not known. L1 expression is extremely low in differentiated cells except for testis (1;2), but it is significantly elevated in breast malignancies (3;4). These observations suggest that posttranscriptional mechanisms are involved in limitation of L1 expression. We demonstrated that the use of the polyA sites located within the L1.3 genome limits the amount of full-length L1.3 mRNAs present in mammalian cells (5). Internal L1 polyA signals can also be functional when fragments of L1.3 elements are inserted into 3' untranslated regions of mammalian genes (our unpublished data). This unique attenuation mechanism helps to minimize the rate of L1 retrotransposition, but may also increase the negative impact of these insertion events on the genome after their insertion. Human EST database searches suggest that these internal polyA signals may also play a role in regulation of L1 expression in a tissue and/or tumor specific manner with breast cancer tissues supporting the least efficient internal L1 polyadenylation. The EST data are strengthened by significant differences in the L1 RNA profiles between breast cancer and non-breast cancer cell lines detected by northern blotting of transiently transfected human L1.3. These observations suggest a potential global change in the mechanism of polyadenylation process upon malignant transformation of mammary gland.

Recent data demonstrate that L1 expression may be toxic to mammalian cells (Gasior and Deininger, under the second review in Nature Genetics). This toxicity is due to introduction of double-stranded breaks into genomic DNA, and it is dependant on L1 endonuclease (EN) domain of ORF2. Our data suggest that L1 toxicity in breast cancer cell lines is not limited to EN activity of ORF2 and it is linked to the apoptotic pathway of cell death.

In summary, we have discovered a novel mechanism that not only attenuates the expression of the only currently active autonomous non-LTR retroelement in mammals, but potentially regulates the production of the full-length L1 mRNA in a tissue specific manner. This discovery has consequences for both cancer initiation and progression. Even a minimal boost in L1 expression via inefficient internal polyadenylation may result in higher levels of L1 related proteins. Increase in L1 EN activity would lead to an increase in DNA damage, which directly contributes to genomic instability and cancer progression.

Body.

According to the approved Statement of Work we have made the following progress in the three years of funding:
Task 1. To identify functional polyA sites in the L1.3 genome. (Accomplished).
Task 2. To identify the function and relative strength of functional polyA sites in normal and breast cancer cells. (Accomplished).
Briefly, we experimentally proved our hypothesis of putative internal polyadenylation signals playing a role in attenuating L1 expression. We confirmed that multiple predicted canonical and noncanonical internal polyA signals identified within L1.3 coding region are functional by site-directed mutagenesis, 3'RACE, and human and mouse EST database analysis. We also demonstrated the attenuation process is redundant, therefore removal of one or even multiple polyadenylation signals from the L1.3 genome results in the more efficient utilization of the polyA sites located in the vicinity (5).

Task 3. To identify the effect of functional polyA sites on the rate of L1 retrotransposition.

To accomplish the final goal of the proposal, we used a wild-type L1.3 expression vector containing a neomycin (Neo^R) selection cassette driven by a SV40 promoter (Fig. 1) (6). The neomycin gene is positioned in the reverse orientation relative to L1.3 transcription. It is interrupted by an intron that can only be spliced following full-length L1.3 mRNA production, but not upon transcription of neomycin gene from the SV40 promoter. Integration of the expression vector directly into cellular DNA would result in production of Neo mRNA interrupted by an intron that cannot be spliced out. Therefore, a functional protein that provides Neo^R can only be produced upon L1.3 mRNA retrotransposition into a new genomic location. These events are selected for by maintaining mammalian cells transiently transfected with the L1.3 Neo plasmid under geneticine selection for at least 14 days. This time is sufficient for HeLa cells to form visible Neo^R colonies. Genocin selection selects for any cell that has Neo^R and as a result it is directly correlates with the rate of L1.3 retrotransposition.

We first optimized transfection and selection conditions for transient transfections of breast relevant cell lines. To mimic L1 retrotransposition potential in normal breast epithelial we used Human Mammary Epithelial (HME) cells. This cell line is immortalized by stable expression of human telomerase, and, therefore, is the best imitation of the normal breast epithelial available to us. MCF7 and Sk-Br-3 cell lines were used to represent malignant mammary cells. Ntera2 and HeLa cell lines were used as non-breast malignancies control.

We compared retrotranspositional potential of the wild-type (wt) L1.3 and 5M L1.3 mutant that lacks five functional polyadenylation (polyA) sites. This mutant of L1.3 has previously been shown to have an increased retrotransposition rate in HeLa cells (5). Retrotransposition assay of wt and 5M L1.3 in HME cells demonstrated an increase similar to the one we observed in HeLa cells. It has been recently demonstrated that elevated L1 expression results in increased number of double-stranded breaks (DSB) in the genomic DNA in HeLa cells (Gasior and Deininger, under the second review in Nature Genetics). Double-stranded breaks are extremely toxic and may lead to cell death. L1 induced DSB

are dependant on L1 endonuclease activity. Our 5M mutant lacks functional polyadenylation signals located in the fist half of the L1.3 genome, which most likely allows some increase in the amount of the full-length L1 mRNA to be made. Even a slight boost in mRNA production would lead to some elevation of L1 proteins and, as a result, L1 endonuclease activity. Therefore, the results of our retrotransposition experiments could be masked by the toxicity of the 5M mutant.

To fairly compare the rate of retrotransposition of the wt and mutant L1.3 elements, we first needed to determine the toxicity level of the wt L1.3 element in different cell lines. In order to accomplish this we also had to address issues associated with potential variations in (i) transfection efficiencies, (ii) promoter activity, and (iii) splicing of the intron interrupting the neo cassette in different cell lines.

To control for transfection efficiency in different cell lines we utilized a pCEP4 vector that was used to clone L1.3 element with the Neo selection cassette for transient transfections of the cell lines of interact. This vector contains Hygromycin B resistance that allows selection for stable integrants into the genomic DNA. The number of Hygromycin B resistant colonies should be proportional to the amount of transfected DNA assuming that there is no variation in the integration step between different cell lines.

Promoter activity of CMV, L1.3; and SV40 promoters that are present in the L1.3 Neo retrotransposition vector was tested by luciferase assay. Each of these promoters was cloned upstream of *Firefly* luciferase, resulting vectors were transiently transfected in HME, SK-Br-3, MCF7, Ntera2, HeLa, and MCF7 stably expressing anti-apoptotic Bcl2 gene. Luciferase analysis demonstrated significant variations in the promoter strength between different cell lines (Fig. 2). We determined that all promoters tested in this study had the highest activity in HeLa cells and the lowest in HME and Sk-Br-3 cell lines. MCF7 and MCFBcl2 cell lines fell in the middle and had similar profiles of the tested promoters.

We have previously demonstrated that splicing of the intron interrupting the Neo^R gene is inefficient (). In NIH 3T3 cells the splicing efficiency of this intron is about 50%. Therefore, only half of the full-length mRNA produced would be retrotranspositionally competent. We utilized northern blot analysis to determine the splicing efficiency in different cell lines (Fig. 3). Our data show that the splicing efficiency varies in different cell lines and therefore it is a contributor to the potential differences in the retrotransposition rates.

Overall our analysis demonstrated that multiple factors may influence the results of retrotransposition assay involving L1 vector tagged with any reporter gene including Neo^R. We have identified two breast cancer cell lines, MCF7 and MCF7 stably expressing Bcl2 that support similar expression of L1.3 and Neo^R gene and similar splicing of the neo intron. These cell lines were chosen to test L1.3Neo toxicity. We used wt L1.3 as well as EN mutant and a double mutant of EN and reverse transcriptase (RT) domains of L1.3. Vectors containing wt and mutant L1.3 elements were transiently transfected in MCF7 and MCF7Bcl2 cells (Fig. 4). The cells were put on Hygromycin B selection for two weeks to select for random integration of the vector.

We did not detect any significant increase in the amount of colonies formed by wt, EN(-), and EN(-)RT(-) L1.3 elements in MCF7 cells. The number of colonies formed by these three vectors in MCF7Bcl2 cell line was about 300-fold higher in the MCF7Bcl2 cells. Hygromycin B selection of cells transfected with a vector that does not express wt L1

element, but carries Hygromycin B resistance (pCEP4) was used as a control. The number of Hyg^R colonies formed by both cell types transfected with pCEP4 plasmid was very similar (data not shown). It was also similar to the number of Hyg^R colonies observed in MCF7Bcl2 cells transfected with wt and mutant L1.3. These results indicate that L1.3 toxicity is not limited to the EN domain of its ORF2 protein, and toxicity levels vary significantly between cell lines. Additionally, our data suggest that presence of the anti-apoptotic protein to a large degree eliminates L1 toxicity in MCF7 breast cancer cell line. A similar experiment was done with wt L1.3 tagged with blasticidin resistance gene (L1.3Blast). This vector was transiently transfected in MCF7 and MCFBcl2 cells and selected with blasticidin for L1 retrotransposition and hygromycin B for random integration (Fig.5). Transient transfection of pCep4 vector in both cell lines was used as a control for random integration in the absence of L1.3 expression. Transfection efficiency and random integration of pCep4 vector was very similar in both cell types. Again we observed a significant increase in the amount of hygromycin B resistant colonies formed by wt L1.3 in MCF7Bcl2 cells. L1 retrotransposition in MCF7Bcl2 cells was also increased by about 20 to 30-fold. Further studies need to be conducted to determine whether this effect is common in other cell lines.

Overall our data demonstrate that L1 integration events represent only a fraction of the damage that L1 expression may impose on the host genome. This discovery has implications for both breast cancer initiation and progression. Therefore delineation of mechanisms by which L1 expression affects the integrity of genomic DNA and cell viability will contribute to our understanding of its role in breast cancer biology.

Overall Key Research Accomplishments.

Task 1. To identify functional polyA sites in L1.3 genome (Months 1-25)
Task 2. To identify the function and relative strength of functional polyA sites in normal and breast cancer cells (Months 25-29).
Task 3. To identify the effect of functional polyA sites on the rate of L1 retrotransposition (Months 29-36)

- We determined that putative canonical and noncanonical polyA sites found in the coding region of the human L1.3 element are functional.
- We demonstrated that removal of five functional polyA signals resulted in an increase in the rate of L1 retrotransposition in HeLa and HME cells.
- We demonstrate by northern blots of transiently transfected breast cancer cells that there is a much less efficient utilization of the internal L1 polyA signals in SK-Br-3 breast cancer cell line.
- We performed human EST database search to determine whether there are any differences in the usage of internal polyA signal in breast tissues.
- We performed human EST database search to determine whether there are any differences in the usage of internal polyA signal between normal human tissues human breast cancers.
- We performed a 3' RACE analysis to determine functional polyA signals in L1 fragment located in the 3' UTR of the reporter gene.
- We determined that L1 toxicity is not limited to its EN activity in MCF7 breast cancer cell line.
- Our data indicate that L1-related toxicity can be overcome by expression of anti-apoptotic protein in MCF7 breast cancer cells.

Reportable outcomes:

In the three-year period we have accomplished the following:

1. An appended manuscript "RNA truncation by premature polyadenylation attenuates human mobile element activity" has been published in journal of *Nature Genetics* Volume 35, Number 4, December 2003.

2. Responses to the publication: Press release "The silence of the LINEs", "Silent garbage"

3. Appended is an abstract "RNA truncation by premature polyadenylation attenuates human mobile element activity" of the presentation that took place at the annual Molecular and Cellular biology Program Retreat, Tulane University, New Orleans, LA in October, 2003.

4. An appended abstract "LINE1 polyadenyletion sites or stability limit elements expression" for the poster that was presented at the Keystone Symposia meeting "Transposition and Other Genome Rearrangements" in Santa Fe, NM, February 8-14, 2003.

4. A Ph.D. in Molecular and Cellular Biology from Tulane University December 2003

5. 17th Morris F. Shaffer and Margaret H.D. Smith-Shaffer Award for Excellence in Research, received in May, 2004.

6. Patent application 60/445,945 (extension filed on February 7, 2004).
Deininger, Prescott L. and Victoria Perepelitsa Belancio
Entitled: Mammalian Retrotransposable Elements.

7. Roy-Engel, A.M., El-Sawy, M., Farooq, L., Odom, G.L., **Perepelitsa-Belancio, V.**, Bruch, H., Oyeniran, O.O., and Deininger, P.L. (2005) Human retroelements may introduce intragenic polyadenylation signals. *Cytogenetic and Genome Res.* (in press).

8. El-Sawy, M., Kale, S., Dugan, C., Nguyen, T.Q., **Perepelitsa-Belancio, V.**, Bruch, H., Roy-Engel, A.M., and Deininger, P.L. (2005) Nickel stimulates genetic instability through L1 retrotransposition. (submitted).

9. Appended is an abstract "The role of polyA tail in the L1 3'end formation" for the FASEB summer research conference "Mammalian Mobile Elements" in Arizona, June 4-9, 2005.

10. Manuscript "The role of polyA tail in the L1 3'end formation" in preparation.

Conclusions:

1. We demonstrated that both canonical and noncanonical polyA sites found in the coding region of the human L1.3 element are functional, and they attenuate the amount of the full-length mRNA and the rate of retrotransposition in HeLa cells.

2. Our EST data and transient transfections of L1.3 expression cassette of human breast cancer cell line suggest that the process of internal polyadenylation may be involved in regulation of the amount of the full-length L1 mRNA in a tissue specific manner and upon malignant transformation.

3. Our 3' RACE analysis demonstrated that internal L1 polyadenylation sites are used during transcription of a luciferase reporter gene when the L1 fragment is cloned into the 3' UTR of the gene. This finding has significant implications for our understanding of mammalian gene expression.

4. L1 toxicity may not be limited to its endonuclease activity. This finding together with finding #2 has implication for L1 related cancer initiation and/or progression.

5. L1-related toxicity can be overcome by expression of anti-apoptotic Bcl2 protein in MCF7 breast cancer cell line.

Reference List

1. Branciforte,D. and Martin,S.L. (1994) Developmental and cell type specificity of LINE-1 expression in mouse testis: implications for transposition. *Mol.Cell Biol.*, **14**, 2584-2592.

2. Bratthauer,G.L. and Fanning,T.G. (1992) Active Line-1 Retrotransposons in Human Testicular Cancer. *Oncogene*, **7**, 507-510.

3. Asch,H.L., Eliacin,E., Fanning,T.G., Connolly,J.L., Bratthauer,G. and Asch,B.B. (1996) Comparative expression of the LINE-1 p40 protein in human breast carcinomas and normal breast tissues. *Oncol.Res.*, **8**, 239-247.

4. Bratthauer,G.L., Cardiff,R.D. and Fanning,T.G. (1994) Expression of LINE-1 retrotransposons in human breast cancer. *Cancer*, **73**, 2333-2336.

5. Perepelitsa-Belancio V. and Deininger P. (2003) RNA truncation by premature polyadenylation attenuates human mobile element activity. *Nat.Genet.*, **35**.

6. Moran,J.V., Holmes,S.E., Naas,T.P., DeBerardinis,R.J., Boeke,J.D. and Kazazian,H.H., Jr. (1996) High frequency retrotransposition in cultured mammalian cells. *Cell*, **87**, 917-927.

7. Hutchison,C.A., III, Hardies,S.C., Loeb,D.D., Shehee,W.R. and and Edgell,M.N. (1989) LINEs and related retroposons: long interspersed repeated sequences in the eucariotic genome. In Berg,D.E.a.H.M.M. (ed.), *Mobile DNA.* American Society for Microbiology, Washington,D.C., pp. 593-617.

8. Nangia-Makker,P., Sarvis,R., Visscher,D.W., Bailey-Penrod,J., Raz,A. and Sarkar,F.H. (1998) Galectin-3 and L1 retrotransposons in human breast carcinomas. *Breast Cancer Res.Treat.*, **49**, 171-183.

Appendices.

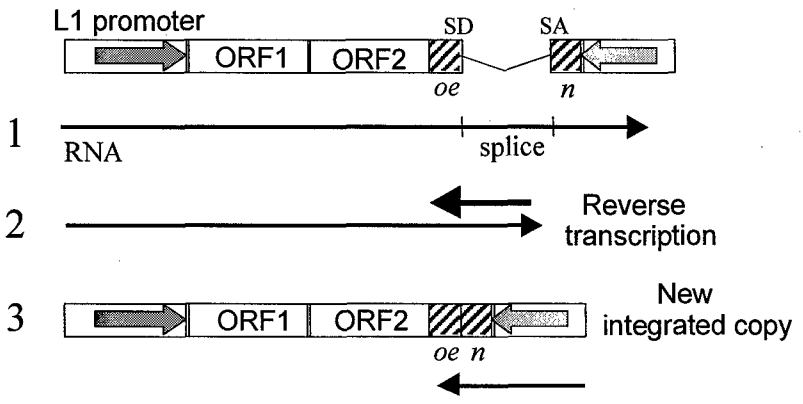


Figure 1. Schematic of the L1 assay system. 1- RNA transcription is performed by the internal L1 promoter. The intron interrupting the neomycin resistance gene is removed by splicing. **2-** The RNA is reverse transcribed, followed by integration of the DNA into the genome. **3-** The new L1 copy contains a neo gene that is functional. Promoters and transcription orientation are indicated by block arrows. SD: splice donor, SA: splice acceptor. The neomycin resistance gene in opposite orientation shown as hatched box. RNA is represented by thin lines with arrows.

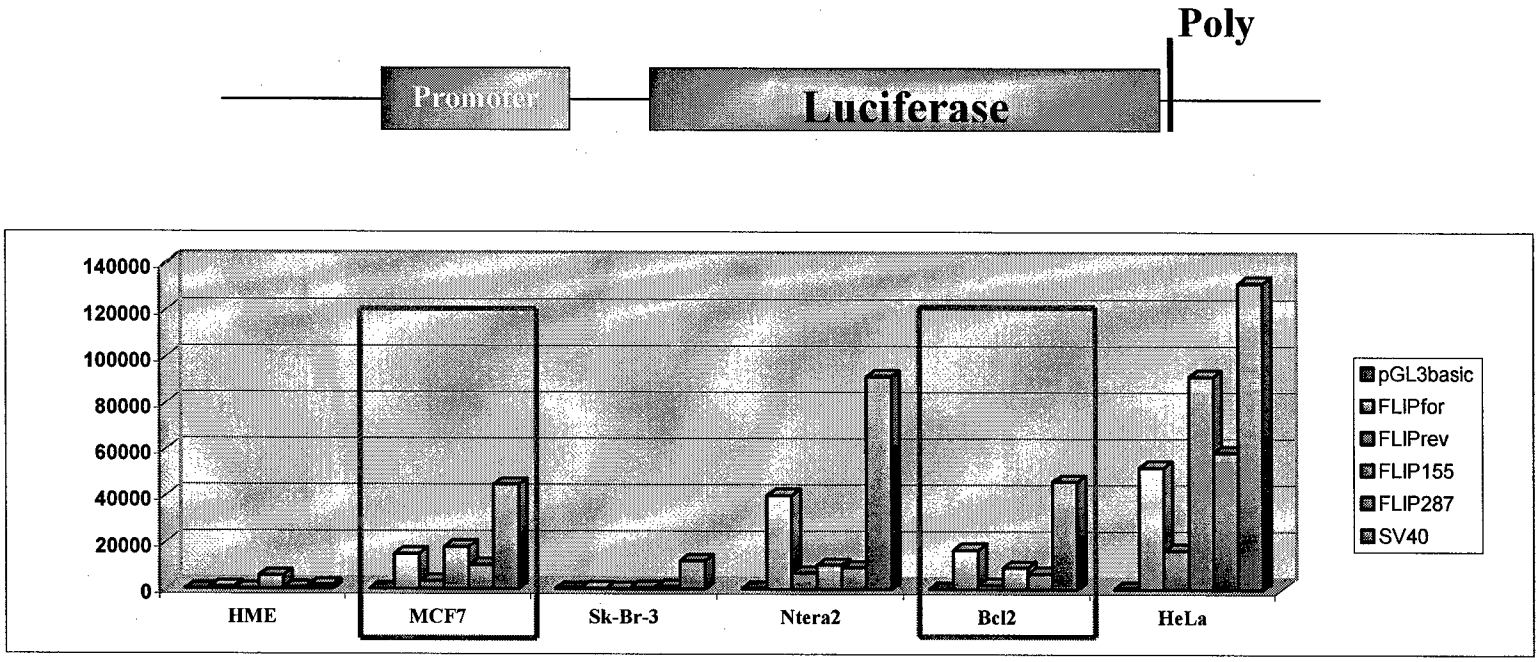


Figure 2. Promoter strength may influence retrotransposition results.
Luciferase expression analysis was performed by transient transfections of a vector containing a *Firefly* luciferase reporter gene driven by FLIP (full-length L1.3 promoter) cloned in either forward (for) or reverse (rev) orientations relative to the reporter gene, FLIP 155 (first 155 bp of FLIP), FLIP 287 (first 287bp of FLIP), and SV40 (Simian virus 40) promoter. PGL3basic (contains no promoter) was used as a negative control. Transient transfections were executed in HME (human mammary epithelial) cells that mimic normal mammary cells, MCF7, Sk-Br-3, MCF7Bcl2 breast cancer cell lines, Ntera2 and HeLa nonbreast cancer cell lines.

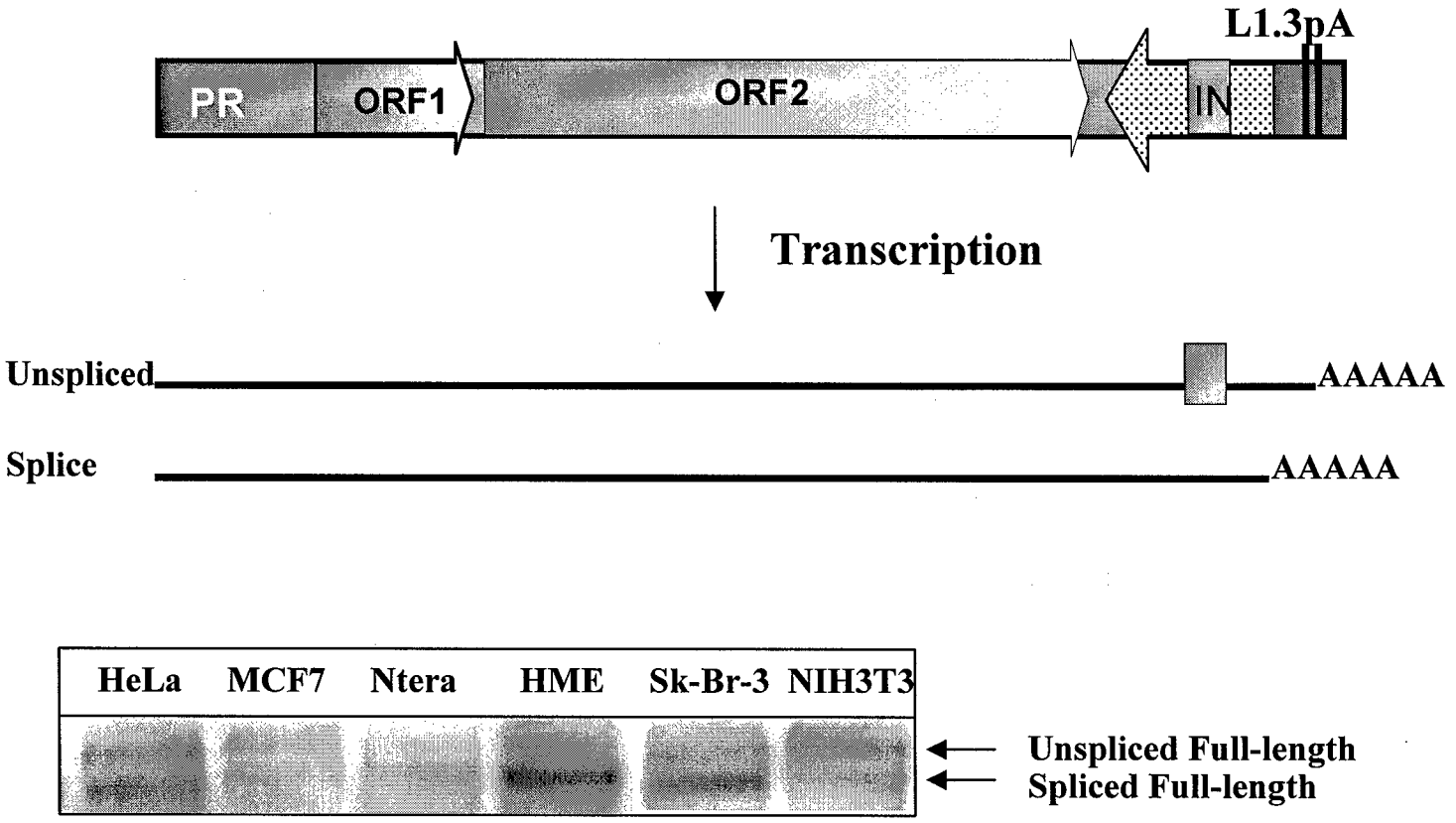
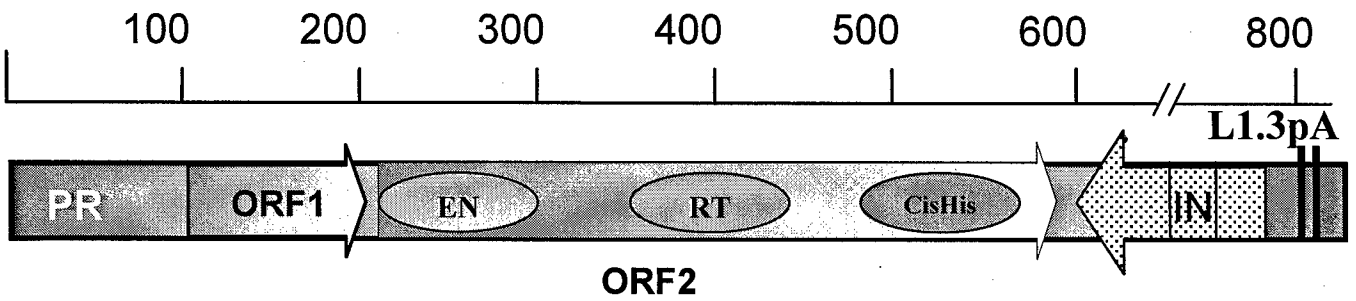
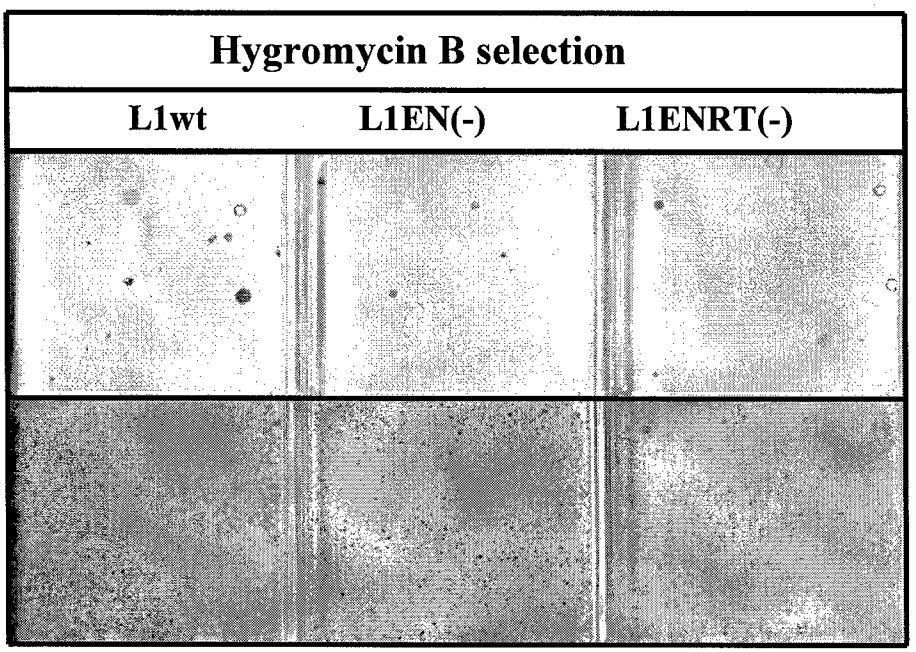


Figure 3. Splicing of L1.3 Neo intron varies between cell types.
Northern blot analysis with strand-specific Neo probe of different cell lines transiently transfected L1.3 containing Neo^R cassette was performed. HeLa, MCF7, Ntera2, HME, and Sk-Br-3 are human malignant and nonmalignant cell lines, NIH 3T3 is a mouse fibroblast cell line.



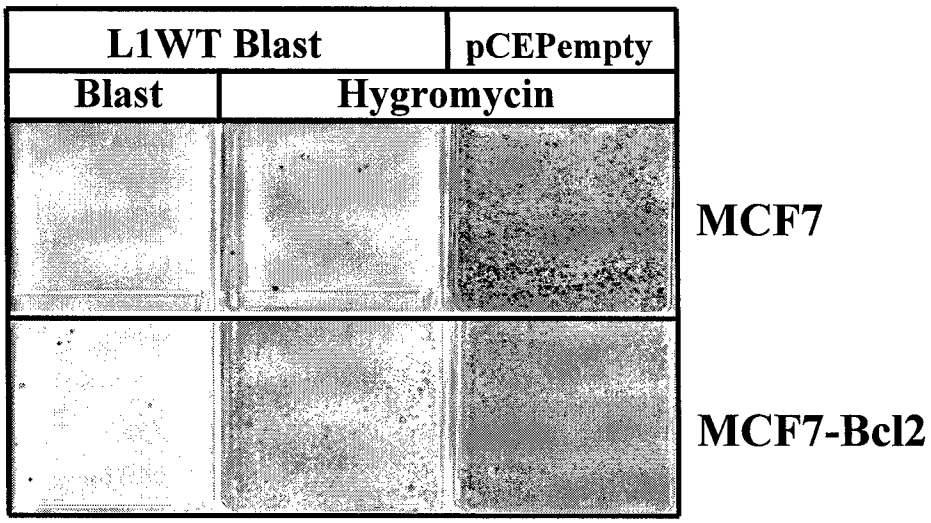


MCF7

MCF7Bcl2

Figure 4. L1-related toxicity is not limited to its endonuclease activity.
MCF7 and MCF7Bcl2 cells were transiently transfected with wild type (L1wt), endonuclease mutant (L1EN(-)), or EN and reversetranscriptase (L1ENRT(-)) mutant and selected with Hygromycin B for random integration of the vector into the host genome.

A.



B.

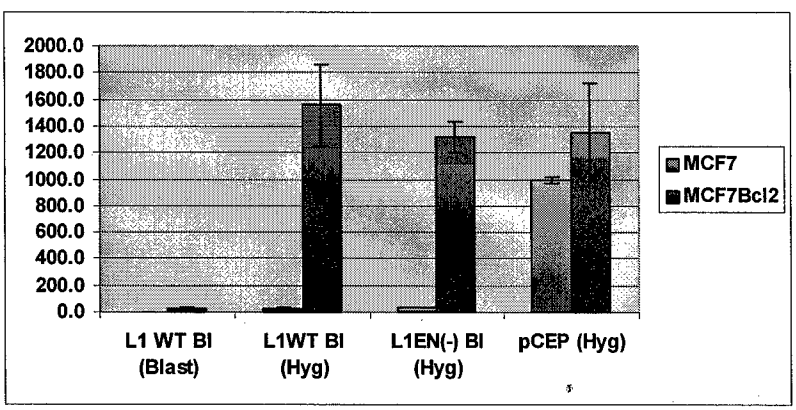


Figure 5. L1 toxicity in MCF7 breast cancer cell line is manifested through Bcl2 apoptotic pathway. Wild type L1.3 with blasticidin selection cassette was transiently transfected in MCF7 and MCF7Bcl2 cells. Blasticidin or hygromycin B selection was used to select for L1 retrotransposition or random integration events respectively. **A.** Visual representation of the resulting colonies. **B.** Quantitative representation of colony formation from multiple experiments.

LINE1 polyadenyletion sites or stability limit elements expression.

Victoria P. Belancio and Dr. Prescott Deininger, Molecular and Cellular Biology,

Tulane University, New Orleans, LA, 70130

LINE1 is the most prevalent human retroelement, and it contributes to genomic instability. LINE1 elements comprise 17% of the human genome, which translates into about $5x10^5$ L1 copies, the majority of which are truncated at their 5' end(7). The full extent of LINE-1 mobility in somatic tissues and particularly in tumors is still not known.
During the course of evolution the activity of LINEs has tremendously decreased. LINE-1 is the only member of the autonomous non-LTR retrotransposones family that is currently active. It is believed that about 60 active copies of full-length LINE-1 are present in the human genome.
L1 expression is extremely low in all cell types of a mature organism except for testis. In contrast, significantly higher levels of LINE-1 expression were found in various cancer cells(4;8). The known factors involved in regulation of LINE1 expression, such as hypomethylation often associated with malignant transformation and promoter activity, cannot fully explain the observed pattern of expression.
In addition, full-length LINE-1 is not detected by Northern blotting in cell culture even when transiently expressed from the CMV promoter. Together, these observations suggest that posttranscriptional mechanisms might be involved in regulation or limitation of L1 expression.
By using a polyadq program, we identified 20 putative polyadenylation (polyA) sites located only in the sense strand of the L1.3 genome. We hypothesize that the use of the putative polyA sites located within the L1.3 genome and RNA instability limit the amount of full-length L1.3 mRNAs present in mammalian cells.

To address this question we employed two biologically relevant to LINE-1 genomic structure systems: (i) bicistronic and (ii) a conventional reporter systems that allow detection of both functional polyA sites and potential RNA instability. We determined by Northern blotting analysis that both of these factors are likely to contribute to a very limited amount of full-length LINE-1 mRNAs present in cells.

RNA TRUNCATION BY PREMATURE POLYADENYLATION ATTENUATES HUMAN MOBILE ELEMENT ACTIVITY.

Victoria P. Belancio and Prescott Deininger

LINE1 is the most prevalent human retroelement, and it contributes to genomic instability. LINE1 elements comprise 17% of the human genome, which translates into about $5x10^5$ L1 copies, the majority of which are truncated at their 5' end(7). The full extent of LINE-1 mobility in somatic tissues and particularly in tumors is still not known. LINE-1 is the only member of the autonomous non-LTR retrotransposones family that is currently active. It is believed that about 60 active copies of full-length LINE-1 are present in the human genome.

L1 expression is extremely low in all cell types of a mature organism except for testis. In contrast, significantly higher levels of LINE-1 expression were found in various cancer cells(4;8). The known factors involved in regulation of LINE1 expression, such as hypomethylation often associated with malignant transformation and promoter activity, cannot fully explain the observed pattern of expression. In addition, full-length LINE-1 is not detected by Northern blotting in cell culture even when transiently expressed from the CMV promoter. Together, these observations suggest that posttranscriptional mechanisms might be involved in regulation or limitation of L1 expression. By using a polyadq program, we identified 20 putative polyadenylation (polyA) sites located only in the sense strand of the L1.3 genome. We hypothesize that the use of the putative polyA sites located within the L1.3 genome and RNA instability limit the amount of full-length L1.3 mRNAs present in mammalian cells.

To address this question we developed a sensitive northern blot assay that allows detection of the full-length L1 mRNA as well as any RNA species produced through internal polyadenylation. We determined that the A-rich coding strand of the L1 elements contains numerous internal polyadenylation sites that attenuate full-length L1 RNA formation by about 50 fold. There is tremendous redundancy in these internal polyadenylation sites, and their presence is conserved throughout mammalian L1 elements. This unique attenuation mechanism helps to minimize the rate of L1 retrotransposition, but may also increase the negative impact of these insertion events on the genome.

POLYA TAIL INVOLVEMENT IN LINE-1 3' END FORMATION.

Victoria P. Belancio, Megan Whelton, and Prescott Deininger.
Department of Epidimiology, Tulane University, New Orleans, LA.

LINE-1 is the only active autonomous non-LTR human retroelement, and it contributes to genomic instability. LINE-1 elements comprise 17% of the human genome, which translates into about $5x10^5$ L1 copies, the majority of which are truncated at their 5' end(7). We have demonstrated that LINE-1 mobility is significantly attenuated by the presence of functional polyadenylation signals throughout the coding region of L1 genomes. A typical polyadenylation site consists of three main cis-acting elements: a conserved hexamer, cleavage site, and a GU-rich downstream region. A newly inserted L1 copy would contain the conserved AATAAA hexamer at the end of its genome. However, the GU-rich downstream region would vary for every L1 copy. The difference in the strength of the polyadenylation (polyA) signal may regulate the efficiency of the L1 3' end formation and, as a result, the amount of the full-length L1 mRNA produced from an individual genomic locus.

A polyA site prediction program, POLYADQ revealed significant variation in the predicted ability of genomic sequences downstream of full-length human L1 elements to support polyadenylation. However, northern blot analysis of the full-length L1 mRNA produced by L1.3 element with genomic sequences predicted to result in either weak or strong polyadenylation signals demonstrated that regardless of the environment the L1 3'end formation was as efficient as in the presence of the SV40 polyA site. Additionaly, the A tail following the conserved AATAAA hexamer positively influences the efficiency of the 3' end in the context of the "weak" downstream sequences, but has minimal effect when adjacent to the "strong" genomic surroundings.

Our study demonstrates that L1 contains sequences that allow efficient 3'end formation upon retrotransposition into a new genomic location independent of base composition downstream of the insertion site. The strategy of 3' end formation by L1 parallels the approach the element employs at its 5'UTR by having an unusual internal pol II promoter.

nature genetics

RNA truncation by premature polyadenylation attenuates human mobile element activity

Victoria Perepelitsa-Belancio & Prescott Deininger

Long interspersed elements (LINE-1s, also called L1s) are the only active members of the autonomous, non–long terminal repeat (LTR) retrotransposon family, which reshapes mammalian genomes in many different ways[1–5]. LINE-1 expression is low in most differentiated cells but high in some cancer cells, in testis and during embryonic development[6–12]. To minimize the negative impact on their hosts' genomes, many mobile elements strategically limit their amplification potential, particularly in somatic cells[13–15]. Here we show that the A-rich coding strand of the human LINE-1 contains multiple functional canonical and noncanonical polyadenylation (poly(A)) signals, resulting in truncation of full-length transcripts by premature polyadenylation. This attenuation lowers the rate of retrotransposition in assays using HeLa cells. It probably also increases the negative effects of LINE-1 insertions into genes[16].

Analysis of the LINE-1.3 sequence with the POLYADQ[17] program identified 19 potential poly(A) signals, AATAAA and ATTAAA, in the sense orientation but only 2 in the antisense strand of the human LINE-1 element (**Supplementary Table 1** online). Several sites are predicted to be much stronger than the relatively weak poly(A) site found at the 3′ end of the LINE-1 element[18]. Gorilla and mouse elements showed a similar enrichment in the sense strand (**Fig. 1**). Additionally, there are 141 noncanonical poly(A) sites distributed throughout the LINE-1.3 genome, deviating from the consensus by only one base (**Supplementary Table 1** online). The widespread presence of these poly(A) signals suggests that they have a conserved function, perhaps limiting, or regulating, LINE-1 retrotransposition.

By transiently expressing an active human LINE-1.3 (ref. 19) element in mouse NIH 3T3 cells, we detected two poly(A), high-molecular-weight, LINE-1.3-specific bands that migrated between 7.4 and 9.4 kb (**Fig. 2a**). The doublet was observed with probes to both the 5′ UTR and the neomycin marker at the 3′ end of the LINE-1.3 construct (**Fig. 2b**), indicating that these RNAs represent full-length transcripts. The presence of the doublet is consistent with the inefficient splicing of the intron[20] (**Fig. 2b**). Multiple, faster-migrating species were also present and much more abundant than the full-length RNAs. The sizes of these transcripts roughly corresponded to the positions expected with use of the putative poly(A) sites identified in the LINE-1.3 sequence (**Fig. 2a**). The most intense band, band 3, correlated with the strongest predicted poly(A) site. The bands were absent from the flow-through fraction of the oligo-dT selection and were not detected by the neomycin strand-specific probe located downstream of these termination sites (data not shown).

In NIH3T3 cells transfected with the LINE-1.3 expression vector lacking the intron-containing *neo*^{r} tag, the pattern of the truncated bands was identical to that detected with the 5′ UTR probe (**Fig. 2a**). Although there was a higher proportion of full-length transcripts, suggesting that the *neo*^{r} tag interferes somewhat with full-length RNA production, these data indicate that the premature polyadenylation is not an artifact of the splicing introduced in the reporter system. We therefore used the LINE-1.3*neo*^{r} expression vector for future experiments to correlate RNA profiles with the rate of retrotransposition.

Northern blotting of RNAs from NIH3T3 cells transfected with a mouse element, LINE-1spa[21,22] (**Fig. 2a**), detected abundant truncated bands corresponding in size to the positions of the putative poly(A) signals in this element. This shows that premature polyadenylation is conserved in mammalian LINE-1 elements. The presence of these polyadenylation sites correlates with the high (~40%) A-residue content in the LINE-1 coding region.

Two inactivating point mutations introduced into the hexamer of the strongest predicted poly(A) site[23] (mutant 1) resulted in the complete loss of the wild-type band 3 in NIH3T3 cells (**Fig. 2a**) and more efficient use of the nearby poly(A) sites (**Fig. 2a**). The amount of full-length LINE-1.3 mRNA (**Fig. 2a**) was similar in the wild type and mutant 1 (**Supplementary Table 2** online). The RNA profile of LINE-1.3 with mutations in canonical poly(A) sites that were expected to produce bands 4 and 5 in the background of mutant 1 (mutant 3) was similar to the band pattern of mutant 1 (**Supplementary Fig. 1** online). This suggested that use of some of the multiple noncanonical poly(A) signals in the region compensates for the removal of the classical sites. Inactivation of the noncanonical poly(A) signals at positions 2,053 or 2,079 in the background of mutant 3 did not significantly change the intensity of band 4 (**Supplementary Fig. 1** online and data not shown). But band 4 disappeared in the mutant lacking both of these poly(A) sites in the mutant 3 background (mutant 5) (**Fig. 2a**). Removal of these poly(A) signals shifted the RNA intensities to higher-molecular-weight bands, indicating more efficient use of the downstream sites, as judged by the intensity of band 2 (**Fig. 2a** and **Supplementary Table 2** online). Thus, internal poly(A) sites present in the LINE-1.3 genome

Tulane Cancer Center, SL66, and Department of Environmental Health Sciences, Tulane University Health Sciences Center, 1430 Tulane Ave., New Orleans, Louisiana 70112, USA. Correspondence should be addressed to P.D. (pdeinin@tulane.edu).

Published online 16 November 2003; doi:10.1038/ng1269

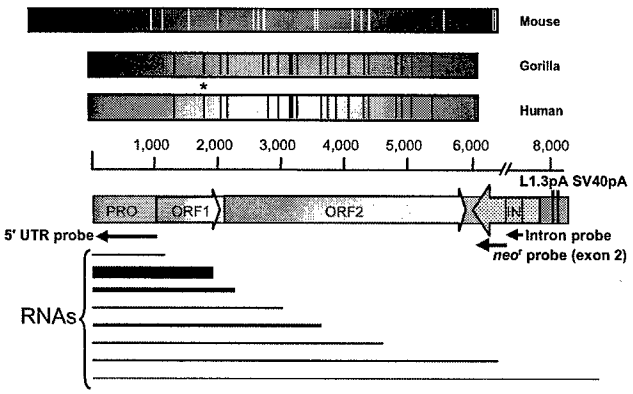


Figure 1 Putative poly(A) sites in the human (LINE1.3), gorilla (LINE-1Gg-1A) and mouse (LINE-1spa) LINE-1 elements. The genomes were aligned according to the beginning of the open reading frame (ORF) 1. An asterisk marks the strongest predicted poly(A) site. Below the nucleotide scale is a schematic of the LINE-1 retrotransposition cassette[29]. The approximate positions of the endogenous LINE-1 promoter (PRO), coding regions (ORF1 and ORF2), 3′ UTR and poly(A) site (L1.3pA), intron (IN) and inverted neomycin resistance gene (large arrow pointing to the left containing the intron sequence) are shown. The SV40 polyadenylation site (SV40pA) is located immediately downstream of the LINE-1 cassette. Horizontal lines represent the predicted mRNA species with the thickness of the line suggesting that different sites probably truncate higher proportions of the RNAs. The bottom dotted line is consistent with read-through transcripts[18]. Thick black horizontal arrows labeled 5′ UTR probe, *neo*^{r} probe and intron probe reflect genomic positions of the strand-specific probes used for northern-blot analysis.

may have a modest effect on limiting the LINE-1.3 expression individually but together result in substantial transcriptional attenuation.

Despite limitations in detection of modest changes of full-length LINE-1 RNA levels, mutants 1 and 5 both caused a statistically significant increase in retrotransposition rate, by a factor of almost 2, in HeLa cells (**Supplementary Table 3** and **Supplementary Methods** online).

Endogenous LINE-1 RNAs from human Ntera2 and HeLa cells had a similar series of truncated transcripts (**Fig. 2c**). We also compared the band distribution in transiently transfected and nontransfected HeLa cells to those in transfected chicken and mouse cells (**Fig. 2c**). We observed similar patterns with only modest differences, suggesting species-specific or cell type–specific variations in use of poly(A) sites. Our results are consistent with a number of previous observations that both human and mouse cells contain a number of smaller, heterogeneous, LINE-1-related bands[6,12,24,25] of unknown origin, with only a few cell types containing full-length LINE-1 RNA[6,12,25]. Full-length transcripts, but not truncated bands, were detected in cytoplasmic, poly(A)$^+$-selected RNAs from Ntera2 cells[25]. Therefore, we tested cytoplasmic versus nuclear poly(A)-selected fractions in transfected NIH3T3 cells and found that the truncated bands, as well as the full-length transcripts, were almost exclusively nuclear (**Supplementary Fig. 2** online).This is consistent with observations that truncated bands from mouse LINE-1 are primarily in the nucleus[12].

We found that most of the transcripts in the human expressed-sequence tag (EST) database with 3′ end sequencing of LINE-1 elements are truncated at positions predicted by the internal poly(A) sites (**Fig. 3**), further confirming the use of the internal poly(A) sites in the endogenous LINE-1 elements. A similar analysis of mouse EST

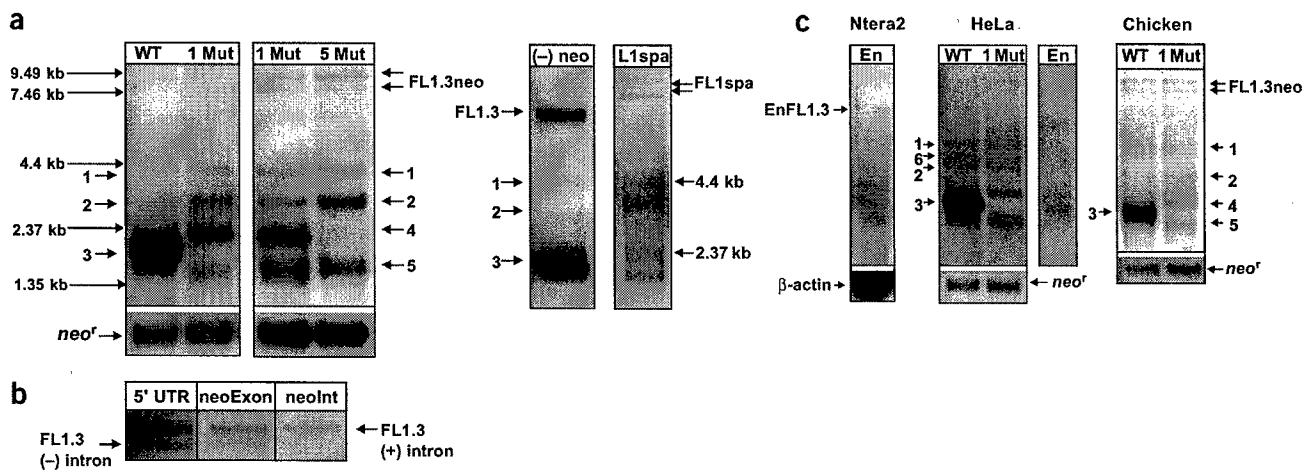


Figure 2 Use of the internal poly(A) sites of LINE-1 element. (**a**) Northern-blot analysis of LINE-1.3 RNA species produced by the wild-type (WT) human element, an element with a mutation in the strongest predicted poly(A) site (mutant 1; 1 Mut), a LINE-1.3 element lacking five poly(A) sites (mutant 5; 5 Mut), wild-type LINE-1.3 element without the *neo*^{r} tag (– neo) and wild-type mouse LINE-1spa (L1spa) in NIH 3T3 cells probed with the human or mouse strand-specific 5′ UTR probe. FL1.3 is full-length human LINE-1.3 mRNA (the size difference of FL1.3 in – neo lane is due to the absence of *neo*^{r} in the expression cassette). FL1spa is full-length mouse LINE-1spa mRNA. *neo*^{r} is neomycin-resistance mRNA detected with the randomly labeled probe to exon 2 of the *neo*^{r} gene. Long horizontal arrows correspond to the positions of the molecular weight RNA marker (Invitrogen). RNA bands that are described specifically in the text are numbered. (**b**) The high-molecular-weight doublet is due to inefficient splicing. A northern blot of LINE-1.3 full-length RNAs produced by the wild-type human LINE-1 element was probed with a strand-specific 5′ UTR probe and probes to either exon 2 (neoExon) or the intron (neoInt) of *neo*^{r}. Bands corresponding to the spliced and unspliced mRNA species are marked FL1.3 (–) intron and FL1.3 (+) intron, respectively. (**c**) Northern blots of endogenous and exogenous LINE-1 RNAs in different species. A northern blot of poly(A)$^+$-selected total LINE-1.3 RNA species produced by the endogenous LINE-1 (EnFL1.3) elements in Ntera2 and HeLa cells and wild-type (WT) and mutant 1 (1 Mut) human LINE-1.3 elements transfected into HeLa and chicken cells was probed with human strand-specific 5′ UTR probe. FL1.3 is full-length human LINE-1.3 mRNA. *neo*^{r} is neomycin-resistance mRNA detected with the randomly labeled probe to exon 2 of the *neo*^{r} gene and was used as a transfection and loading control. β-actin is β-actin mRNA detected by randomly labeled probe to *ACTB* and was used as loading control for endogenous mRNAs in Ntera2 cells.

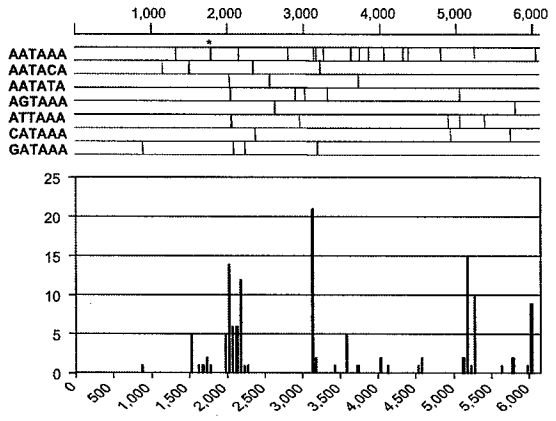


Figure 3 3′ ends of ESTs relative to common LINE-1 poly(A) signals. Map positions are shown for the strongest predicted canonical and noncanonical poly(A) sites in the LINE-1.3 sequence. The asterisk indicates the strongest poly(A) site in our transfection experiments. The chart below shows the relative abundance of human ESTs whose 3′ ends map approximately to the locations shown. This represents 140 ESTs that had been oligo-dT primed and sequenced from the 3′ end. There are only eight full-length LINE-1 elements in this group, representing less than 6% of the total LINE-1-related transcripts. The largest cluster resides around position 2,000, in the region seen near our strongest predicted poly(A) site, as well as those shown in the mutant of that site, and the positions seen in the northern blot of endogenous LINE-1 elements from Ntera2 cells.

(**Supplementary Fig. 3** online) supported the conservation of the process of internal polyadenylation of LINE-1 RNA in mammals. There were few ESTs corresponding to the position of band 3 (**Fig. 2a**) for human LINE-1, but more in the regions of bands 4 and 5 (**Fig. 2a**), consistent with our observations that the relative use of the internal poly(A) sites varied in different cell types. Band 3 may be particularly strong in the cultured cells used in these studies, whereas other sites may be preferred in the endogenous tissues used for most EST studies.

Chicken fibroblasts that lack any endogenous LINE-1 elements but support use of LINE-1 poly(A) signals were transfected with LINE-1.3 (**Fig. 2c**). We used 3′ RACE to amplify the region of the primary truncated LINE-1 RNA products in the northern blots. This analysis showed a pattern consistent with the northern blot results for the wild type and mutant 1 (**Fig. 4a**). There were also multiple larger bands in both elements that corresponded in length to the LINE-1.3 region containing poly(A) signals responsible for the band 4 in the northern-blot assay. This confirms the sites of the functional poly(A) signals and that the relative activity of any given site varies depending on competition with other nearby poly(A) signals[26]. Sequence analysis of eight clones from the strongest 3′ RACE band (**Fig. 4a**) confirmed the use of that site. Analysis of 13 clones from the 2,053–2,153 region of the 3′ RACE gel (**Fig. 4a**) identified 7 clones that terminated at a location consistent with the use of the ATTAAA poly(A) site that was inactivated in mutant 5 (**Fig. 2a**), 3 clones that terminated at a location consistent with the use of the third AATAAA site in LINE-1 (**Fig. 1**), which was inactivated in mutant 3, and 3 clones that used various noncanonical sites located in that general region of LINE-1. All polyadenylations occurred at sites consistent with distance and with the sequence normally associated with 3′ cleavage[26–28]. We carried out 3′ RACE amplification of poly(A)$^+$-selected mRNAs from nontransfected and wild-type LINE-1.3-transfected Ntera2 cells (**Fig. 4b**). The bands were similar in size to those identified in chicken cells. Band 1,788, corresponding to the position of the strongest poly(A) signal, was not detected in the nontransfected Ntera2 cells. It was observed on transfection with the LINE-1.3 expression cassette, however, consistent with differences between endogenous elements in the EST data (**Fig. 3**) and our northern blots of LINE-1.3-transfected cells (**Fig. 2**).

We showed that LINE-1 sequences support the use of a broad range of poly(A) sites that limit retroposition activity of the element, resulting in less retroposition and, therefore, less damage to their hosts' genomes. Despite the low levels of full-length LINE-1 RNA, LINE-1 retrotransposition rates can be quite high, suggesting that the later steps of L1 integration may be very efficient. On the other hand, the internal poly(A) sites may increase the potential damage when LINE-1 elements insert into a gene, causing premature termination of that gene's transcripts. Because of the redundancy in the system, elimination of one or two poly(A) sites may effect only minimally the overall amplification capability of the element. Our observation of a new application of the conventional mRNA 3′ end formation extends the range of the known mechanisms for the control of mRNA levels in mammalian cells and provides additional understanding of the selective pressures that act on new retrotransposition events in reshaping the mammalian genome.

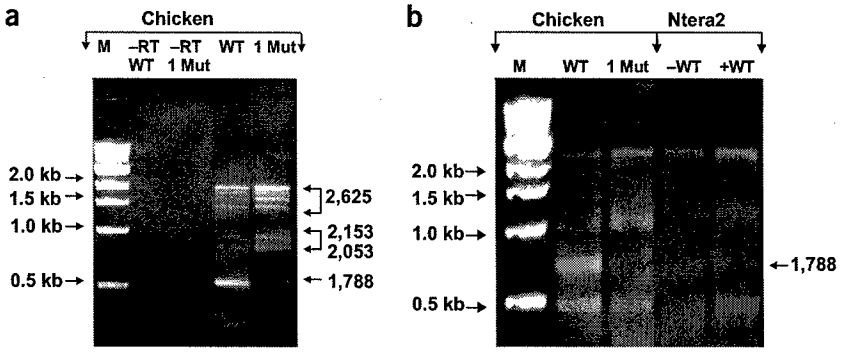


Figure 4 3′ RACE analysis of the prematurely terminated LINE-1.3 RNA species. (**a**) 3′ RACE analysis of poly(A)-selected total mRNAs from chicken fibroblasts transiently transfected with the wild-type (WT) or mutant 1 (1 Mut) LINE-1.3 expression vectors described in **Figure 1**. The upstream primer used in the PCR step corresponds to positions 1,342–1,359 of the LINE-1.3 sense strand. M, 1-kb DNA ladder (NEBiolabs); –RT, without reverse transcriptase. The band corresponding to the strongest predicted poly(A) site identified by the northern blot is labeled 1,788. Genomic location and the hexanucleotide sequence of the identified functional poly(A) sites just upstream of individual bands, after cloning and sequencing, are shown next to their positions on the gel. (**b**) 3′ RACE analysis of poly(A)-selected total mRNAs from chicken fibroblasts transiently transfected with the wild-type (WT) or mutant 1 (1 Mut) LINE-1.3 expression vectors described in **Figure 1** and Ntera2 cells nontransfected (–WT) and transfected with the wild-type LINE-1.3 expression cassette (+WT). The upstream primer used in the PCR step corresponds to positions 1,159–1,188 in the LINE-1.3 sense strand.

METHODS

Cell culture. We maintained NIH 3T3 (ATCC CRL-1658), HeLa (ATCC CCL2) and Ntera2 (ATCC CRL-1973) cells at 37 °C and 5% CO_2 in Dulbecco's modified Eagle medium (GIBCO) high glucose, 10% Colorado calf serum (GIBCO); minimal essential medium (GIBCO), 10% fetal bovine serum (GIBCO); and Dulbecco's modified Eagle medium high glucose (GIBCO), 15% fetal bovine serum, respectively. We maintained chicken fibroblasts (ATCC CRL-12203) at 39 °C and 5% CO_2 in Dulbecco's modified Eagle medium high glucose (GIBCO) and 15% fetal bovine serum.

Transient transfection assay. We transfected 4×10^6 NIH 3T3 cells and 6×10^6 HeLa cells, Ntera2 cells or chicken fibroblasts per 75 cm^2 cell culture flask (Corning) with 3–6 μg of the LINE-1.3 expression cassette by lipofectamine (6–12 μl of Plus reagent, 18–36 μl of lipofectamine; Invitrogen) 16–18 h after plating. We incubated NIH 3T3 cells with the transfection cocktail in the serum-free medium for 4 h and incubated chicken fibroblasts, HeLa cells and Ntera2 cells with the transfection cocktail in the serum-free medium for 3 h. We incubated all cell types in their respective media for 24–26 h before collecting RNA.

RNA extraction and poly(A) selection. We combined the contents of four 75-cm^2 cell culture flasks of each cell type and extracted total mRNA using TRIzol Reagent (Invitrogen). We then carried out chloroform extraction and isopropanol precipitation. We used the PolyATract mRNA isolation system III (Promega) to select poly(A) RNA species as instructed by the manufacturer. We resuspended poly(A)-selected and precipitated RNA in 30 μl of RNase-free water and fractionated it in a single lane of an agarose-formaldehyde gel.

Northern blots. We transferred RNA to a Hybond-N nylon membrane (Amersham Pharmacia Biotech) by capillary transfer overnight at room temperature in a standard 5× sodium chloride/sodium citrate (SSC) solution. We crosslinked the RNA to the membrane with ultraviolet light and prehybridized it in 30% formamide, 1× Denhardt's solution, 1% SDS, 1 M NaCl, 100 μg ml^{-1} salmon sperm DNA and 100 μg ml^{-1} yeast tRNA at 60 °C for at least 6 h. Hybridization with a strand-specific probe (final concentration of $4–8 \times 10^6$ cpm ml^{-1}) was carried out overnight in the same solution at 60 °C. We carried out multiple 10-min washes at high stringency (0.1× SSC, 0.1% SDS) at 60 °C. We quantified the results of the northern-blot assays on a Fuji Phosphorimager.

We generated the strand-specific probe used for the northern-blot assay by the MAXIscript T7 system (Ambion). Primer sequences for generating the template are available on request. We produced DNA template for the probe by PCR with primers that amplified either the LINE-1.3 5′ UTR or exon 2 or the intron of the *neo*^{r} cassette. The T7 promoter sequence was included in the reverse primer of each pair. We fractionated PCR products on a 1% low-melting agarose gel, excised them and purified them using a QIAquick gel extraction kit 50 (QIAGEN).

3′ RACE analysis. We carried out 3′ RACE (Clontech) according to the manufacturer's protocol on poly(A)-selected total mRNA from chicken fibroblasts and Ntera2 cells transiently transfected with LINE-1.3 expression cassette as described above. Primer sequences are available on request. We gel-purified DNA with QIAquick Gel Extraction Kit (QIAGEN), cloned it into pCR2.1-TOPO vector (Invitrogen) and sequenced 3–8 clones for each poly(A) site by automated sequencing.

Site-directed mutagenesis. We used the QuikChange Site-Directed Mutagenesis kit (STRATAGENE) to change the AATAAA hexamer sequence of the strongest predicted poly(A) site in the LINE-1.3 genome (position 1,788) as well as the poly(A) sites at positions 1,323 and 2,154; the poly(A) sites at position 1,323 and 2,154 to GATCAA; the ATTAAA hexamer sequence of the poly(A) signal at position 2,053 to ATCAAG; and the GATAAA hexamer sequence of the poly(A) signal at position 2,079 to GATCAA.

EST searches. We identified LINE-1 ESTs using a BLAST search of the entire human EST database with the LINE-1.3 sequence. We collected the first 1,085 hits, but only 140 corresponded to 3′ sequences. Those sequenced from the 5′ end would not necessarily define the poly(A) site. For **Figure 3**, the 3′ positions of the ESTs were binned into 50-base regions.

Accession numbers. LINE-1.3, L19088; LINE-1Gg-1a, AF036235; LINE-1spa, AF016099.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank C. Schmid and M. Batzer for helpful comments and J. Moran, H. Kazazian Jr. and J. Goodier for providing LINE-1 vectors and for discussion. This work was supported by the US National Institutes of Health (P.D.) and the US Department of Defense Breast Cancer Research Program (V.P.-B.).

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 10 September; accepted 27 October 2003
Published online at http://www.nature.com/naturegenetics/

1. Lander, E.S. *et al.* Initial sequencing and analysis of the human genome. International Human Genome Sequencing Consortium. *Nature* **409**, 860–921 (2001).
2. Deininger, P.L., Moran, J.V., Batzer, M.A. & Kazazian, H.H. Jr. Mobile elements and genome evolution. *Curr. Opin. Genet. Dev.* **13**, 1–8 (2003).
3. Ostertag, E.M. & Kazazian, H.H. Jr. Biology of L1 retrotransposons. *Annu. Rev. Genet.* **35**, 501–538 (2001).
4. Kazazian, H.H. Jr. & Moran, J.V. The impact of L1 retrotransposons on the human genome. *Nat. Genet.* **19**, 19–24 (1998).
5. Esnault, C., Maestre, J. & Heidmann, T. Human LINE retrotransposons generate processed pseudogenes. *Nat. Genet.* **24**, 363–367 (2000).
6. Branciforte, D. & Martin, S.L. Developmental and cell type specificity of LINE-1 expression in mouse testis: implications for transposition. *Mol. Cell Biol.* **14**, 2584–2592 (1994).
7. Trelogan, S.A. & Martin, S.L. Tightly regulated, developmentally specific expression of the first open reading frame from LINE-1 during mouse embryogenesis. *Proc. Natl. Acad. Sci. USA* **92**, 1520–1524 (1995).
8. Ostertag, E.M. *et al.* A mouse model of human L1 retrotransposition. *Nat. Genet.* **32**, 655–660 (2002).
9. Bratthauer, G.L., Cardiff, R.D. & Fanning, T.G. Expression of LINE-1 retrotransposons in human breast cancer. *Cancer* **73**, 2333–2336 (1994).
10. Asch, H.L. *et al.* Comparative expression of the LINE-1 p40 protein in human breast carcinomas and normal breast tissues. *Oncol. Res.* **8**, 239–247 (1996).
11. Skowronski, J., Fanning, T.G. & Singer, M.F. Unit-length line-1 transcripts in human teratocarcinoma cells. *Mol. Cell. Biol.* **8**, 1385–1397 (1988).
12. Benihoud, K. *et al.* Unusual expression of LINE-1 transposable element in the MRL autoimmune lymphoproliferative syndrome-prone strain. *Oncogene* **21**, 5593–5600 (2002).
13. Deininger, P.L. & Batzer, M.A. Mammalian retroelements. *Genome Res.* **12**, 1455–1465 (2002).
14. Lerat, E., Capy, P. & Biemont, C. Codon usage by transposable elements and their host genes in five species. *J. Mol. Evol.* **54**, 625–637 (2002).
15. Siebel, C.W. & Rio, D.C. Regulated splicing of the Drosophila P transposable element third intron in vitro: somatic repression. *Science* **248**, 1200–1208 (1990).
16. Medstrand, P., van de Lagemaat, L.N. & Mager, D.L. Retroelement distributions in the human genome: variations associated with age and proximity to genes. *Genome Res.* **12**, 1483–1495 (2002).
17. Tabaska, J.E. & Zhang, M.Q. Detection of polyadenylation signals in human DNA sequences. *Gene* **231**, 77–86 (1999).
18. Moran, J.V., DeBerardinis, R.J. & Kazazian, H.H. Jr. Exon shuffling by L1 retrotransposition. *Science* **283**, 1530–1534 (1999).
19. Dombroski, B.A., Mathias, S.L., Nanthakumar, E., Scott, A.F. & Kazazian, H.H. Isolation of an active human transposable element. *Science* **254**, 1805–1810 (1991).
20. Gilbert, N., Lutz-Prigge, S. & Moran, J.V. Genomic deletions created upon LINE-1 retrotransposition. *Cell* **110**, 315–325 (2002).
21. Mulhardt, C. *et al.* The spastic mouse: aberrant splicing of glycine receptor beta subunit mRNA caused by intronic insertion of L1 element. *Neuron* **13**, 1003–1015 (1994).
22. Naas, T.P. *et al.* An actively retrotransposing, novel subfamily of mouse L1 elements. *EMBO J.* **17**, 590–597 (1998).
23. Sheets, M.D., Ogg, S.C., & Wickens, M.P. Point mutations in AAUAAA and the poly (A) addition site: effects on the accuracy and efficiency of cleavage and polyadenylation in vitro. *Nucleic Acids Res.* **18**, 5799–5805 (1990).
24. Martin, S.L. Ribonucleoprotein particles with LINE-1 RNA in mouse embryonal carcinoma cells. *Mol. Cell. Biol.* **11**, 4804–4807 (1991).
25. Skowronski, J. & Singer, M.F. Expression of a cytoplasmic LINE-1 transcript is regulated in a human teratocarcinoma cell line. *Proc. Natl. Acad. Sci. USA* **82**, 6050–6054 (1985).
26. Zhao, J., Hyman, L. & Moore, C. Formation of mRNA 3′ ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol. Mol. Biol. Rev.* **63**, 405–445 (1999).
27. Colgan, D.F. & Manley, J.L. Mechanism and regulation of mRNA polyadenylation. *Genes Dev.* **11**, 2755–2766 (1997).
28. Proudfoot, N. Poly(A) signals. *Cell* **64**, 671–674 (1991).
29. Moran, J.V. *et al.* High frequency retrotransposition in cultured mammalian cells. *Cell* **87**, 917–927 (1996).